DTU Library

Observations and modeling of the diurnal SST cycle in the North and Baltic Seas

Karagali, Ioanna; Høyer, J.L.

Published in:
Journal of Geophysical Research - Oceans

Link to article, DOI:
10.1002/jgrc.20320

Publication date:
2013

Observations and modeling of the diurnal SST cycle in the North and Baltic Seas

I. Karagali[1], and J.L. Høyer[2]

I. Karagali, DTU Wind Energy, Technical University of Denmark, Risø Campus, Building 125, Roskilde, 4000, DK. (ioka@dtu.dk)

J.L. Høyer, Centre for Ocean and Ice (COI), Danish Meteorological Institute (DMI), Lyngbyvej 100, 2100, DK. (jlh@dmi.dk)

[1]DTU Wind Energy, Risø Campus, Roskilde, 4000, DK.

[2]COI, DMI, Copenhagen Ø, 2100, DK.

Abstract. This paper discusses the evaluation of three parameterizations for the diurnal variability of SST during one year, from February 2009 to January 2010 (inclusive), using high resolution, regional atmospheric model outputs as input fields. Comparison of the spatial extent of diurnal warming in the Northern European Seas from SEVIRI and the models, indicates the ability of the models to reproduce the general patterns seen from the observations. Mean absolute biases between the SEVIRI observed peak warming and the modeled results do not exceed 0.25 K, with a maximum standard deviation of 0.76 and a 0.45 correlation. When random noise is added to the models, their ability to reproduce the statistical properties of the SEVIRI observations improves. The correlation between the observed and modeled anomalies and different parameters highlights the importance of wind as a driving field. A positive correlation is found between hourly SEVIRI anomalies and the daily mean diffuse attenuation coefficient $K_{d(490)}$.

1. Introduction

Properly resolving the daily cycle of sea surface temperature (SST) is of scientific interest due to its many implications. Accounting for the daily SST variability is considered a priority to be implemented in climate and numerical weather prediction (NWP) models as errors in the estimated fluxes can be large [*Webster et al.*, 1996; *Ward*, 2006]. Blended SST fields are commonly used as boundary conditions to initialize models. These blended products are representative of foundation temperatures, i.e. when the water column is well mixed and no diurnal signal is present. But the diurnal variability must be known when blending SST fields from different satellite sensors with different overpass times. Including a diurnal variability field in the operationally available daily SST fields would simplify the generation of the blended products. Hourly satellite observations and in situ observations when available, are valuable for the quantitative and qualitative description of the daily SST cycle. Nonetheless, their implementation in real time forecasting is complicated due to the potential computational cost, the missing values in satellite SSTs that would require some kind of interpolation and the infrequent and potentially spurious in situ measurements.

Many attempts to develop models for the diurnal variability of SST have been made and their complexity varies from empirical parameterizations to bulk mixed layer models and turbulent closure models. An extensive review is available from *Kawai & Wada* [2007]. *Price et al.* [1986] used a data set collected during the Long-Term Upper Ocean Study (LOTUS) to examine diurnal warming. A mixed layer model was developed, relating warming with the surface wind stress and surface heating. It was found to have some success in simulating the amplitude and day to day variability of the diurnal cycle.

Fairall et al. [1996] developed a warm layer model based on a single-layer scaled version of the *Price et al.* [1986] model. When combining it with their cool skin model and testing against data obtained from the Tropical Ocean-Global Atmosphere (TOGA) Coupled Ocean-Atmosphere Response Experiment (COARE) program, they reported a model underestimation of 0.5 K for the observed 3–4 K warming amplitude and a phase lag of 1 hour.

Zeng & Beljaars [2005] derived a prognostic scheme for the computation of skin SST, required in applications such as weather forecasting and coupled ocean-atmosphere modeling. They compared modeled results with radiometric measurements from a ship over the Western Pacific warm pool region. The mean absolute deviation between computed and observed skin temperatures was 0.39 K and the correlation was 0.85. The averaged diurnal cycle over a 10-day period had an observed amplitude of 2.3 K while the modeled amplitude was 2 K.

Gentemann et al. [2009] developed the Profiles of Ocean Surface Heating (POSH) model, to determine the vertical profile of surface heating, based on the model described in *Fairall et al.* [1996] and additional measurements from the M-AERI and SkinDeEP instruments. Comparison of 72 days with measured diurnal warming showed a mean bias of -0.01 K and a standard deviation of 0.28 for modeled minus measured warming.

Observations from the TOGA-COARE program were also used by *Webster et al.* [1996] to examine the amplitude of the peak daily warming. Based on model simulations using a one-dimensional mixed layer model, a parametrization was developed. This used as input parameters the peak solar insolation, the 10-m wind speed and the daily averaged precipitation. The reported overall bias was lower than 0.05℃, maximized at 0.07℃ for very high wind speeds.

Kawai & Kawamura [2002] used a one dimensional model to simulate skin and 1-m depth temperatures. These were used for the derivation of a regression equation to evaluate the daily amplitude of SST from daily mean wind and daily peak solar radiation. They found greater skin SST values than in *Webster et al.* [1996] but not inconsistent with observations, as they stated.

Gentemann et al. [2003] calculated an empirical model using non-linear least squares regression, to relate the daily SST variability from Pathfinder and Tropical Microwave Imager (TMI) datasets to insolation and wind speed. The initial daytime minus night-time differences were 0.22°C for the bias with a 0.68°C standard deviation. When the modeled diurnal warming was subtracted from the daytime PathFinder SSTs, the reported difference dropped to −0.04°C with a standard deviation of 0.63°C.

Clayson & Curry [1996] used the model results from *Webster et al.* [1996] to determine the amplitude of the diurnal cycle and developed a parametrization of peak daily warming. Measured peak warming was found to be 0.13 ℃ higher than the one modeled using ship-measured input parameters, with a standard deviation of 0.31 ℃ and a correlation of 0.85. *Clayson & Weitlich* [2007] used the parametrization of *Webster et al.* [1996] to produce modeled SST from satellite input fields. Comparisons with buoy data showed a mean bias of 0.0012 ℃, a standard deviation of 0.26 ℃ and a correlation of 0.74.

Most of the models mentioned above were derived using in situ observations and are based on regression analysis. Lately, *Filipiak et al.* [2012] developed a model relating the diurnal variability of SST with modeled net surface heat flux and surface wind speed fields. They adopted an approach different than in previous studies, and attempted to ensure that the model predicts correctly the statistical distribution of diurnal warming.

Diurnal warming in the Northern European Shelf Seas from hourly SEVIRI observations has been formerly identified and characterized in *Karagali et al.* [2012]. Using five years of hourly SST retrievals (2004-2009), they showed that anomalies higher than 2 K occurred during the late spring and summer months of every year. *Merchant et al.* [2008] noted influential variability of the diffuse attenuation coefficient at 490 nm, $K_{d(490)}$, in the North Sea and Baltic Sea that could explain up to 20% of the mean peak dSST spatial variance.

In the present paper, different parameterizations that predict diurnal warming are evaluated and compared to the SEVIRI observed day-time anomalies in the Northern European Seas. Observations from this area of the world have not been generally considered when developing the diurnal warming models, therefore it is of interest to evaluate their performance in an area where diurnal variability has only recently been characterized and quantified. As such, it is also sought to define the correlation between the observed warming with the modeled anomalies and the fields from a regional atmospheric model. Finally, as the Northern European Seas are characterized by turbid waters due to the outflow of large rivers, the potential impact of the mean attenuation coefficient $K_{d(490)}$ on the development of diurnal warming cases is examined.

The model from *Filipiak et al.* [2012], the prognostic scheme of *Zeng & Beljaars* [2005] and the parametrization from *Clayson & Curry* [1996] are implemented. The selection is based on their simple implementation, low computational cost and different modeling approaches. Observations from the Northern European Seas have been included in the

Filipiak et al. [2012] model, therefore it is of interest to evaluate the model's performance when different input fields are used.

The data are presented in Section 2 and the models are briefly described in Section 3. Section 4 contains results separated in two parts. Initially, SEVIRI and modeled warming cases are compared in terms of spatial extent and statistics. The latter part describes sensitivity analyses and the correlation of observed anomalies to various physical parameters and the modeled warming. A discussion can be found in Section 5 and the conclusions in Section 6.

2. Data

2.1. Satellite Data

The Spinning Enhanced Visible Infrared Imager (SEVIRI), on board the Meteosat Second Generation (MSG) satellites, is an infra-red radiometer. Radiation is collected from an area using a telescope and is focused on detectors sensitive to 12 bands of the electromagnetic spectrum [*Aminou et al.*, 1997]. The nadir sampling distance is 3 km for the near infra-red and infra-red channels.

The experimental SEVIRI product of CMS (Centre Météorologie Spatiale), Météo France is used from February 2009 to January 2010 (inclusive). It is different from the operational EUMETSAT OSI SAF product as it is released for every hour and is mapped to a 0.05 degrees grid. A flag ranging from 0 (unprocessed) to 5 (excellent) indicates the quality of the SST values, with 3 being acceptable. SEVIRI SSTs are adjusted to sub-skin temperatures.

A Surface Solar Irradiance (SSI) experimental product is also available from CMS, Météo France for the O&SI SAF, EUMETSAT. It is calculated as the solar irradiance reaching the Earth's surface in the 0.3–4 μm band, with the irradiance being defined as the radiant flux per unit area (W m^{-2}). A full description of the algorithms and the processing chain is available in *O&SI SAF* [2005]. The product used here is hourly, available on a grid of 0.1° resolution. For the present study, data are re-sampled at 0.05°, to match the SEVIRI grid.

The diffuse attenuation coefficient at 490 nm (K$_(d490)$) is a measure of the optical properties of the water column. Higher $K_{d(490)}$ values indicate higher water turbidity and tend to promote diurnal warming [*Merchant et al.*, 2008]. For the purposes of the present study, fields from two different sources are obtained.

The GlobColour project (`http://www.globcolour.info/`) produces daily $K_{d(490)}$ fields based on the merged chlorophyll-A concentration for case 1 waters [*GlobColour*, 2010]. The $K_{d(490)}$ product from the Danish Meteorological Institute (DMI), is produced by an algorithm developed specifically for Case-II waters during the REgional VAlidation of Meris chlorophyll Product (REVAMP) project [*Peters et al.*, 2005]. Due to its regional character, it does not cover the entire domain, but extends only from 5°W to 26°E.

2.2. NWP fields

The NWP High Resolution Limited Area Model (HIRLAM) is developed by several European meteorological institutes and has a spatial resolution of 3 km [*Mahura et al.*, 2005]. HIRLAM outputs are used to obtain 10 m winds, the surface net heat flux, the solar and non-solar fractions of the surface fluxes and the averaged daily precipitation. All fields are re-sampled to match the SEVIRI grid.

3. DV Models

Filipiak et al. [2012] Model

This model, here referred to as FMKLB, evaluates warming from dawn to next dawn. It has been derived using 1 year of data over the Atlantic and the Mediterranean Sea from SEVIRI SST observations and operational and analysis fields from the European Centre for Medium-Range Weather Forecasts (ECMWF). Diurnal warming (D) is described as a function of time (t) and maximum wind speed (W) and integrated net heat flux (Q) since the net heat flux (q) becomes positive. The warming part is described by

$$D\left(t\right) = Q\left(t\right)\frac{\alpha\left(t\right)}{1 + b\left(t\right)W_t{}^2} + c\left(t\right), \tag{1}$$

where α, b, c are derived coefficients that depend on the hour of the day. Cooling periods are defined by negative integrated net heat flux (Q$<$ 0) and are described as

$$D\left(t\right) = f\left(t\right)\frac{Q\left(t\right)}{\rho c_p d}, \tag{2}$$

where ρ is the density and c_p, the specific heat of water, d is the climatological mixed layer depth and f is a coefficient that defines the fraction of d that undergoes cooling. For the climatological mixed layer depth d, the entire water column depth is used in this study, as both the North Sea and the Baltic Sea are characterized by very small depths.

Zeng & Beljaars [2005] Model

The prognostic scheme of *Zeng & Beljaars* [2005], referred to as ZB, requires wind and surface fluxes, along with an a priori knowledge of foundation temperatures. SST$_{found}$ fields composed from multi-day, night-time SEVIRI SSTs of quality 3 and higher are used, as representative of night-time conditions [*Karagali et al.*, 2012]. The scheme, as described in *Zeng & Beljaars* [2005] consists of two parts. The sea surface skin temperature T_s is described as

$$T_s - T_{-\delta} = \frac{\delta}{\rho_w c_w k_w} \left(Q + R_s f_s\right), \tag{3}$$

where δ is the thickness of the skin layer, $T_{-\delta}$ is the sub-skin temperature, ρ_w is the density and c_w the volumetric heat capacity of sea water, k_w is the molecular thermal conductivity of sea water, R_s is the net solar radiation at the surface of the ocean and f_s the fraction absorbed in the sub-layer. Q represents the sum of the surface sensible and latent heat fluxes and the net long-wave radiation. Below the skin layer the temperature is described by

$$\frac{\vartheta \left(T_{-\delta} - T_{-d}\right)}{\vartheta t} = \frac{Q + R_s - R\left(-d\right)}{d \rho_w c_w \nu / \left(\nu + 1\right)} - \frac{\left(\nu + 1\right) k u_{*w}}{d \varphi_t \left(d/L\right)} \left(T_{-\delta} - T_{-d}\right), \tag{4}$$

where d is the depth where no diurnal warming occurs, T_{-d} is the foundation temperature, φ_t is the stability function, u_{*w} is the friction velocity in the water and ν is an empirical parameter considered equal to 0.3 for a d=3 m. In the present study, two different thresholds are selected, d_1=3 m and d_2=6 m. In addition, as the scheme computes both the skin and sub-skin SST, we use the sub-skin to be comparable with SEVIRI.

Clayson & Curry **[1996] Model**

The parametrization, referred to as CC, was developed within the framework of the TOGA-COARE project and it is described originally in *Webster et al.* [1996]. This scheme uses a peak solar insolation, the averaged daily wind speed and the daily average precipitation through a regression equation of the form

$$DSST = \alpha + b\,(PS) + c\,(P) + dln\,(U) + e\,(PS)\,(U) + f\,(U)\,. \tag{5}$$

$DSST$ is the peak skin SST, PS is the peak solar insolation in W m^{-2}, U is the mean daily wind speed in m s^{-1}, P is the averaged daily precipitation in mm hr^{-1} and a, b, c, d, e, f are coefficients derived through regression and are determined separately for $U \geq 2$ m s^{-1} and $U < 2$ m s^{-1}. For the remainder of this paper, the term dSST is used to describe the warming at any given time of the day while dSST$_{max}$ is the maximum value of dSST.

The CC scheme also requires a foundation temperature field, for which the SEVIRI SST$_{found}$ are used as in the ZB scheme. In the present study, a Neural Network (NN) system, developed by Bogdanoff & Clayson (personal communication), is used for the determination of coefficients and the DSST prediction. From the above, it follows that the CC scheme estimates the skin SST and does not resolve the full daily cycle but only produces a peak value. Therefore, comparison with the other models and SEVIRI can be achieved only through the distribution of dSST$_{max}$ and the mean dSST$_{max}$ and it will include the "cool skin" bias.

4. Observing and modeling the diurnal cycle

An example of the day-time dSST$_{max}$ from SEVIRI, the FMKLB, ZB and CC models during a warming event on the 03/07/2009, is shown in Figure 1. SEVIRI dSST$_{max}$ amplitudes up to 5.5 K are observed in an extended area off the west coast of Denmark and smaller areas, where dSST$_{max}$ is ~4 K, are observed in the waters around the country. Another area with dSST$_{max}$ of 5 K is observed in the Baltic Sea. Each model shows a rather different behavior, with the FMKLB capturing warming in the Danish inland waters and in the Baltic Sea but failing off the west coast of Denmark. The ZB d_1 scheme predicts dSST$_{max}$ of the correct amplitude but with a much larger spatial extent while the d_2 version essentially reproduces the same spatial extent but with a reduced amplitude. Finally, the CC scheme shows a much lower dSST$_{max}$, on the order of 2 K; this scheme predicts the skin SST and thus, the potential large difference between skin and sub-skin temperatures is highlighted.

Visual inspection of the early morning wind fields from NASA's QuikSCAT and ESA's ENVISAT ASAR, taken at ∼06:00 and ∼10:00 correspondingly, showed winds lower than 4 m s^{-1}. In addition, the morning retrieval from AVHRR (∼10:00) was used to compute the day-time anomalies from the SEVIRI SST$_{found}$ field. AVHRR showed warming of 1 K for extended areas and cases with warming up to 5 K were identified. The spatial structure of warming had similarities with the SEVIRI observed dSST$_{max}$ of Figure 1 but the large event off the west coast of Denmark was not identifiable yet.

Figure 2 shows the HIRLAM mean wind speed and Q from 10:00 to 16:00 and the GlobColour K$_d$(490) field on the 03/07/2009, similar to the diurnal warming example shown previously. All modeled warming scenarios show a clear dependence on the HIRLAM wind field. The integrated net heat flux is high for the largest part of the domain. K$_d$(490) values are generally low but some spatial variability can be identified with values around the west coast off of Denmark and the Danish waters being higher.

4.1. Statistical Description

To evaluate the performance of the models, their statistical distributions are compared to the ones derived from SEVIRI. However, noise is inherent in the SEVIRI results, defined as the difference between the foundation and day-time SST fields while all systematic noise is ignored. *Karagali et al.* [2012] reported a minimum standard deviation of ∼0.3 K between the SEVIRI SST_{found} fields and validation fields consisting of the last pre-dawn SST quality 5 value. When considering the difference between the day-time SST fields and the SST_{found}, this standard deviation is expected to be higher. In addition, they also reported a standard deviation of ∼0.8 K between SEVIRI SST_{found} and in situ measurements from moored buoys. This standard deviation includes noise on time scales larger than a day. Based on these minimum and maximum standard deviation estimates, the random noise is defined as white noise with zero mean and a standard deviation of 0.5 K. The DV models estimate the "true" diurnal cycle without noise. The random noise component can have a significant impact in the shape of the probability density function (PDF) and for consistency, we show the PDFs with (white) and without (gray) the addition of random noise. The percentiles and sums of data points are estimated only with the added noise.

The distribution of anomalies equal to or larger than 2 K is shown in Figure 3, for SEVIRI and the different models. In particular, the FMKLB scheme reproduces the distribution shown from SEVIRI, especially without the added random noise. The model's behavior does vary with the added noise; an increase of anomalies in the range 2–2.3 K and a decrease for the range 2.4–3.5 K when compared to the results with no random noise, is observed. The number of identified anomalies increases when the random noise is added. The FMKLB percentiles are ∼0.2 K lower than the SEVIRI ones. The FMKLB model has been developed using 1 year of SEVIRI observations, thus the noise inherent in SEVIRI is expected to be also inherent in the model. This may explain why the distribution is more similar to SEVIRI without the added noise.

The distribution of the ZB d_1 version does not change when adding the random noise, but the number of identified warming events increases. The distribution is similar to SEVIRI for dSST$\geq$2.5 K while the 95th and 97.5th percentiles are higher than the SEVIRI ones. Modeled anomalies do not exceed 5 K in amplitude while the 99.9th percentile from SEVIRI is at 6.25 K. The highest impact of adding random noise is seen for the ZB d_2 version. Without adding noise, no dSST higher than 3 K is identified. With the noise, the distribution resembles more the one from SEVIRI but the percentiles are significantly lower.

The distribution of time of warming is shown in Figure 4. SEVIRI warming cases are evenly distributed during the day, with a peak at 15:00 LT and warming that persists until 22:00 LT. Adding random noise significantly improves all models. Especially, the FMKLB anomalies peak an hour earlier, at 14:00 and the noise adds warming until 22:00; originally no warming was observed after 20:00. The ZB scheme shows a time lag of 1 hour for the d_2 version. The random noise significantly improves the left part of the distributions as warming is identified earlier, which is also seen from SEVIRI. There is an overestimation of late warming from both ZB versions, which is slightly reduced when the random noise is added.

The distribution of the duration of warming is presented in Figure 5. All distributions are scaled to the highest 99.9 percentile; this is observed from SEVIRI and is estimated at 13 hours. The 99.9 percentiles for the models are 13 hours for FMKLB, 12 hours from ZB d_1 and 10 hours for ZB d_2. SEVIRI has the 75th percentile at 3 hours. Adding the noise significantly modifies the model distributions towards the SEVIRI one. The FMKLB and ZB d_2 have the same percentiles as SEVIRI. The ZB d_1 version has higher percentiles suggesting that warming is not sufficiently dissipated. This can be an artifact of the shallow warm layer setting (d_1 =3 m).

Figure 6 shows the distribution of dSST$_{max}$, estimated as the daily maximum anomaly of every grid cell, independent of threshold. The CC model is also included for comparison as it can only produce a daily skin dSST$_{max}$. Only anomalies with quality flag 5 are considered for SEVIRI. 75% of the SEVIRI daily dSST$_{max}$ do not exceed 1.1 K but 5% are higher than 2.4 K. Note the relatively low number of total observations, due to missing values either in the day-time SSTs or the night-time reference fields.

The distribution of dSST$_{max}$ from the FMKLB model has the 75th percentile higher by $\sim$ 0.2 K but its higher percentiles are lower than SEVIRI. The ZB d_1 scheme has the 75%, 95% and 97.5% of dSST$_{max}$ higher than SEVIRI; only the 99.9% is lower by $\sim$0.7 K. The d_2 version produces a distribution of dSST$_{max}$ with percentiles following the behavior of the FMKLB model. Finally, the CC scheme has a much narrower distribution with no dSST$_{max}$ exceeding 2 K and 75% not exceeding 0.7 K. Adding random noise modifies the distributions for all the models, resulting in closer fits to the SEVIRI distribution. Especially the shape of the distribution for the noise-modified CC model is closest to the SEVIRI for warming up to 1–1.5 K. Thus, depending on which statistical properties are of interest, the FMKLB scheme may be considered more appropriate as it resolves better the magnitude of extreme cases while the CC scheme performs well when the mean warming is of interest.

The number of dSST$_{max}$ is different between the schemes and SEVIRI, especially between the FMKLB and the rest. This is because both the ZB and the CC require a foundation temperature field, for which the SEVIRI SST$_{found}$ are used. Thus, missing values in the reference fields will prevent the estimation of an anomaly value while this is not the case for the FMKLB scheme; this only uses NWP fields that do not have missing values. In addition, missing SST values from the daytime SEVIRI fields or the SST$_{found}$ fields will result in fewer identified anomalies for SEVIRI.

The impact of the varying amount of identified dSST$_{max}$, between the models, on the shape of the distributions has been examined by filtering the FMKLB dSST$_{max}$ based on the availability of the ZB dSST$_{max}$. Results not shown here indicate that even when the total number of FMKLB dSST$_{max}$ is reduced to half of the original amount, which is shown in Figure 6b, no significant variation of its statistical properties occurs. Percentiles increase by 0.1 K, the noise modified distribution remains practically unchanged while the original distribution shows a more even spread of dSST$_{max}$ in the range 0–0.4 K compared to the one shown in Figure 6b.

The averaged monthly diurnal cycles from the FMKLB and ZB schemes and the SEVIRI observations are presented for the period April–July 2009 in Figure 7. The mean diurnal cycle is evaluated hourly from all grid cells with an anomaly. No random noise is added to the models and the quality of the SEVIRI anomaly must be 5. The cycles are estimated separately for SEVIRI and each of the schemes, therefore they represent the average diurnal signal without any requirement of spatial and temporal collocation.

With the transition from spring to summer warming starts earlier; SEVIRI shows anomalies at 08:00 in April and at 07:00 in May while warming starts between 04:00

and 05:00 in June and July. $dSST_{max}$ reaches up to 0.75 K and lasts longer in summer compared to spring. There is a residual warming of 0.2–0.3 K at night, consistently observable during all months. The coldest part of the cycle is observed between 04:00 and 05:00 and it is negative in April and May indicating that the night-time hourly SEVIRI values are colder than the SST_{found}. In June and July, the coldest part of the cycle is observed around similar hours but it is as cold as the SST_{found}.

The FMKLB model reproduces the time of $dSST_{max}$ but underestimates the amplitude. It also shows slower cooling and does not dissipate at night-time. This late-evening residual warming from the FMKLB model is almost the same as from SEVIRI for April, higher by approximately 0.1 K for May and June and lower in July. The warm layer is destroyed during the early morning hours, as the one from SEVIRI. The FMKLB scheme has a minimum at 04:00 LST, since it is computed from dawn to the next dawn. The existence of the residual warming and its decreasing trend is a positive attribute but there is a significant mismatch in the cooling rate and magnitude when compared to SEVIRI.

The ZB scheme, independent of the d value, shows an early initiation of warming compared to SEVIRI in April and May. The peak amplitude is generally underestimated by at least 0.2 K, except in June when both versions show an underestimation of less than 0.05 K compared to SEVIRI. The ZB scheme generally peaks earlier and cools faster than SEVIRI. It does maintain the late-night residual warming, which does not exceed 0.15 K and is always lower than that observed by SEVIRI. To the contrary, the scheme shows no early morning cooling. The addition of random noise to the models does not modify the monthly shape of the diurnal cycle.

Using only quality 5 daytime SSTs from SEVIRI may bias the monthly cycles to higher estimates. When warming from SEVIRI is computed using values flagged with quality 3–5 (red dashed line), there is an average shift of ∼0.1 K towards lower values when compared to the quality 5 cycle. This shift occurs mostly during the late-night and early-morning cooling parts but in July it occurs throughout the entire day. The peak amplitude difference between the two cycles is smaller than 0.03 K in April and June, zero in May and approximately 0.1 K in July. The warming rate is higher from April to June for the flagged 3–5 cycle, which also peaks earlier in April. Cooling occurs generally faster than the quality 5 cycle and it may be related to sub-pixel clouds. Especially after the peak amplitude, such clouds may be generated locally due to small scale convection and their sub-pixel size can not be identified by the cloud detection scheme [*Le Borgne et al.*, 2012].

Analyses of the monthly averaged diurnal cycles were also performed using a 2 K threshold (not shown). Any grid cell exceeding this threshold at least once from 08:00–22:00 was used with no spatial collocation requirements. The averaged diurnal cycle for SEVIRI was well described by the ZB d_1 version in terms of the peak amplitude and its timing. The FMKLB scheme slightly underestimated the peak amplitude and peaked approximately 2 hours earlier. The ZB d_2 version slightly underestimated the peak and its timing was 2 hours later but best approximated the late-evening warm layer seen from SEVIRI. The quality 5 cycle was approximately 0.1 K higher than the quality 3–5 cycle during the afternoon cooling (April–July) and early-morning cooling and warming (June and July).

4.2. Spatial Distribution

The spatial distribution of hours with warming of 2 K or more, for 1 year of SEVIRI observations and model runs is shown in Figure 8. Note the different scaling to make the spatial differences visible. All models capture the difference in the spatial extent of warming between the North Sea and the Baltic Sea; more warming is identified in the Baltic Sea similar to the SEVIRI results. Good agreement between the modeled and observed anomalies is found for the Baltic Sea around 60°N, the North Sea around 58°N, 4°E and the Danish Straits around 56°N, 12°E. For an area west of Denmark (54°N, 8°E), SEVIRI warming events are observed but the schemes do not succeed in resolving it. The spatial correlation between SEVIRI and the models is estimated based on the grid cells with warming $\geq$2 K and the number of hours and is shown in Table 1. The highest correlation is between SEVIRI and the ZB d_1 version, which nonetheless does not exceed the 0.5 value.

When random noise is added to the models (not shown), the spatial distribution of warming exceeding 2 K is slightly extended. An increase of approximately 20 to 40 hours is observed in areas where a significant number of warming events is already identified with no random noise. The FMKLB scheme shows the lowest increase, $\sim$ 20 hours, while the ZB d_2 version shows the largest increase, $\sim$ 40 hours. In most areas where no warming is observed without noise, adding noise produces a few hours of warming greater than 2 K; nonetheless, their number ranges between 1 and 10 for all the models. The noise-modified spatial correlation with SEVIRI, shown in Table 1, increases marginally.

The NWP input fields used in the parameterizations control the concurrence of favorable conditions for diurnal warming. Figure 9 shows the number of cases for which u is equal to or lower than 5 m s^{-1} and at the same time the surface net heat flux q is higher than 300 W m^{-2}. In general, high coincidences of conditions that promote diurnal warming are found in areas where the parameterizations do predict most anomalies, i.e. the southern part of the English Channel, the coastal areas of the Baltic Sea, the Irish Sea and off the west coast of Norway. The area around 54°N, 8°E where SEVIRI anomalies have been identified but the models do not perform accordingly, appears in Figure 9 as an area with a relatively low number of coincident low u and high q events.

Averaged dSST$_{max}$ is shown in Figure 10. The scaling is different for the FMKLB and CC models to highlight the spatial variability. SEVIRI averaged dSST$_{max}$ is around 0.3–0.6 K for a large part of the study area. dSST$_{max}$ higher than 0.6 K is observed off the west coast of Denmark, the Baltic Sea, the west coast of Norway and the Irish Sea. The models capture the dSST$_{max}$ in some of these areas;off the west coast of Norway (58°N–60°N, ∼6°E), the Baltic Sea (54°N–60°N, ∼20°E–30°E) and the Irish Sea (52°N–54°N, ∼4°W). They also show a similar pattern as SEVIRI, where more warming is identified in the Baltic Sea compared to the North Sea. In general, higher dSST$_{max}$ is observed from SEVIRI compared to the models.

The statistics of observed−modeled dSST$_{max}$ are shown in Table 2, for the cases when a value is available from SEVIRI and the models. All coincident dSST$_{max}$ values for each grid cell and every day are used, without adding any random noise. Mean biases are almost zero for the FMKLB and ZB schemes and ∼0.25 K for the CC scheme, highlighting the skin vs. sub-skin difference. From the tested hourly parameterizations, the FMKLB has the lowest bias and standard deviation. The computed correlation coefficients have almost zero "p" values and are, therefore, considered statistically significant.

4.3. Sensitivity Analyses

To investigate the relation between observations, model outputs and different forcing parameters, we compute the correlation between the input fields used in the parameterizations (the integrated net heat flux Q and the wind U), the SEVIRI observed dSSTs and other physical parameters (the attenuation coefficient $K_{d(490)}$ and the sea surface solar irradiance (SSI) from SEVIRI), that are thought to promote warming. This is performed for SEVIRI anomalies exceeding a threshold of 0.5 K. One would expect a negative correlation between u and warming while positive correlations would be expected between warming and $K_{d(490)}$, Q, SSI and its integrated value (SSI'). Integration for the net heat flux and SSI is performed from 03:00 LST until the time of observed SEVIRI dSST. All estimated correlation coefficients have an almost zero “p” value and always much lower than 0.05 therefore are considered statistically significant.

SEVIRI anomalies of 0.5 K and above, occurring between 13:00 and 16:00 LST, are collocated with the corresponding wind from HIRLAM, SSI from SEVIRI, $K_{d(490)}$ from DMI and the GlobColour project and dSST from the FMKLB and ZB models. Note that both $K_{d(490)}$ products are daily, thus only 1 value is available for every grid cell. To overcome this and proceed with the sensitivity analysis, the same value of the grid cell is used for all the collocations. Figure 11 shows the correlation matrix, with the numbers of the actual r values superimposed. On average, the correlation between SEVIRI dSSTs and all other included parameters shows the correct trends but weak signals.

Note the negative but low (-0.32) correlation with the wind speed U and the positive but very small correlation with the $K_{d(490)}$ products. The three schemes show a relatively high ($\sim$0.7) and negative correlation with U, which is one of their forcing fields. Their correlation with the integrated heat flux Q is positive but rather low ($\sim$0.3) while the one with $K_{d(490)}$, is negative and greater in absolute value compared to the corresponding SEVIRI–$K_{d(490)}$. These results are derived without any added random noise. If the noise is added to the models, their correlation with SEVIRI and the various parameters decreases. The two different $K_{d(490)}$ products are well correlated (0.7) despite the fact that they are for case I and II waters. The high correlation between Q and SSI' (∼0.87) is considered a positive result for the validation of Q used in the models, without ignoring the fact that Q includes more heat terms (sensible, latent, longwave).

The time-lagged correlation of different parameters with the SEVIRI dSST is also estimated from 3 hours prior to 3 hours after the time of the SEVIRI anomaly. Results not shown here indicate that the highest (negative) correlation with U is observed for the hours prior to the time of the anomaly. Such is the case for the SSI but for the integrated SSI' and net heat flux Q, the highest correlation is found for the time of warming (and later). When the different parameterizations are considered, there is no time lag in the correlation and the ZB scheme is the one with the highest value (∼0.42).

The correlation values found from this analysis are relatively low but the correct trends are captured. When the threshold on acceptable warming cases is increased from 0.5 to 2 K (not shown), SEVIRI dSSTs are slightly more correlated with the $K_{d(490)}$ products. Nonetheless, the correlation between SEVIRI dSST and model warming decreases and so does the correlation between SEVIRI dSST and U, SSI and Q. Such findings indicate that when filtering only according to SEVIRI dSST values, the spatial mismatch of warming between SEVIRI and the models, and consequently between SEVIRI and the NWP fields (U and Q), gives rise to the low correlation values. Even more so when the threshold on warming increases, in which case the models do not properly resolve the peak amplitude of warming. Regarding the decrease of correlation with SSI when the threshold of warming increases, it is justified by the fact that warming typically occurs with a time lag of some hours in respect to solar noon. Thus, for high warming values, typically occurring around 15 LST, SSI is already decreasing due to the time of day.

5. Discussion

Of the three parametrization schemes applied, only two can resolve the daily cycle. Both tend to properly identify the different warming patterns between the Baltic Sea and the North Sea. They slightly overestimate the spatial extent of warming, with ZB d_1 doing more so than the rest. From a large number of coincident dSST_{max} values the bias between SEVIRI and the models is not larger than 0.25 K and the highest standard deviation does not exceed 0.76 K.

Filipiak et al. [2012] compared their model with SEVIRI observations for a period that was not used during the model development and stated that the mean warming matched to 0.05 K. *Zeng & Beljaars* [2005] found a mean absolute deviation of 0.39 K and a correlation of 0.89 with in situ radiometric measurements for the skin SST. *Clayson & Curry* [1996] used ship measurements for the peak solar radiation, precipitation and wind to derive the skin dSST_{max} and compared it with ship measurements, concluding that the derived warming was lower by 0.13°C with a standard deviation of 0.31°C and a correlation of 0.85.

The mean biases shown in Table 2 are not far from their findings, especially for the FMKLB and ZB schemes. This study has revealed higher standard deviations than the ones in *Zeng & Beljaars* [2005] and *Clayson & Curry* [1996], while the correlation values reported here are almost one half of those reported above. It should be noted that the biases reported in *Zeng & Beljaars* [2005] are for the skin SST while in this study the comparisons are with sub-skin SST. The CC bias, defined in this study as sub-skin minus skin, is almost double in absolute value and of the opposite sign compared to the one in *Clayson & Curry* [1996], where it is estimated using skin measurements.

As such, the biases mentioned in the above studies and the ones given here are not directly comparable. Ideally, in situ observations from drifting buoys would be most suitable for the validation of the SEVIRI and modeled warming. Unfortunately, in the North Sea and Baltic Sea only moored buoys and platforms are available, with typically the shallowest sensor located no less than 1 m below the surface. Examination of such moored buoy observations showed that almost no diurnal cycle could be identified for the period 02/09–01/10, thus eliminating the possibility for validation with in situ measurements.

Successful modeling of the diurnal cycle when using empirical parameterizations critically depends on how well the coefficients represent the local conditions. When NWP model outputs are used as input fields, their spatial and temporal accuracy in resolving the conditions that promote diurnal warming also controls the successful modeling of the diurnal cycle. When evaluating the number of observations where u <=5 m s^{-1} from HIRLAM and other available model outputs, it is found that for the same time period HIRLAM generally predicts low winds more often. In the case that HIRLAM is biased towards low winds, warming will be more often predicted by the parametrization schemes.

Nonetheless, comparisons of HIRLAM minus in situ data were performed for the area where the mismatch between SEVIRI and modeled warming is identified from Figure 8 and where fewer occurrences of low winds and high surface net heat flux are found in Figure 9. In situ measurements of 10 m winds from a meteorological mast at Horns Rev, located at 55°31'N, 07°47'E, about ∼20 km off the west coast of Denmark were compared to HIRLAM winds. From 160 pairs, the HIRLAM minus in situ mean bias was -0.12 m s^{-1} and the root mean square (rms) error 1.1 m s^{-1}. When only low winds $\leq$6 m s^{-1} were examined, 95 pairs gave a mean bias of 0.082 m s^{-1} and an rms error of 1.2 m s^{-1}. This in situ validation is based on very few data points but it does not show that HIRLAM overestimates low winds at this particular location. In addition, internal DMI reports indicate that HIRLAM winds are biased high in the North Sea (Amstrup, personal communication, March 15 2013).

The above findings are inconclusive and rather contradictory. From the available information, HIRLAM may be biased towards lower winds in some areas like the Baltic Sea but may be biased toward higher winds in the North Sea. Unfortunately, extensive validation studies of HIRLAM winds with in situ observations are not available and are beyond the scope of this study. Nonetheless, as also indicated by the correlation of the modeled dSST with wind, which is much higher than with the integrated net heat flux Q, the parameterizations are more sensitive to the input wind fields. Thus, if low winds are predicted more often than they occur, the parameterizations will overestimate warming while they will underestimate warming in the event that the modeled winds are biased high.

In addition, these parameterizations do not account for advection caused by surface currents and tidal motion. The ocean surface warming, identified through SEVIRI SST fields, can be caused locally due to the combination of low enough winds and strong solar heating but it may also be a result of advected water masses and SST gradients. Hence, spatial mismatches between SEVIRI and modeled warming may also be a product of such physical phenomena. Also, the test period of one year may be not representative regarding the performance of the models. A longer time period for the comparisons would eliminate the possibility of particular conditions occurring for the test year that the HIRLAM model failed to capture.

When examining the spatial extent of either individual warming cases (Fig. 1 for the FMKLB, ZB $d = 6$ and CC cases), or of the mean dSST$_{max}$ (Fig.10 for all models) or of anomalies $\geq$2 K (Fig.8 for FMKLB and both ZB schemes), a consistent feature is the lack of modeled warming off the west coast of Denmark in contrast to the SEVIRI observable warming. This area of the North Sea poses challenges due to its very shallow waters, intense tidal effects and high water turbidity. Low overall SEVIRI minus in situ biases from moored buoys and platforms in this area were reported in *Karagali et al.* [2012], slightly increasing during daytime. Such results lower the possibility of regional SEVIRI biases in this complex area but the lack of drifting buoys does not allow for more accurate validation in shallow layers.

The coefficients used in the parametrization of CC are evaluated through a Neural Network (NN) which includes multiple already derived coefficients. Based on the input fields for the solar insolation, the wind and the precipitation the NN predicts the skin dSST$_{max}$. Thus, the implementation of the CC parametrization is straightforward and with a low computational cost. Nonetheless, skin dSST$_{max}$ is expected to be lower than the SEVIRI derived dSST.

As stated in *Filipiak et al.* [2012], their model's coefficients should be re-fitted if NWP fields other than the ECMWF ones are used. In this study we used HIRLAM NWP fields without modifying the coefficients. The ECMWF wind fields used for the model derivation had a 6-hour time resolution, while HIRLAM winds are available hourly. These differences in the type of field, their time and spatial resolution may be a reason for the mismatch between modeled and observed warming. Despite this, the FMKLB model successfully reproduces the statistical distribution of anomalies and some properties of the averaged diurnal cycle seen from the SEVIRI data.

The FMKLB scheme was also evaluated using the climatological mixed layer depth from *de Boyer Montégut et al.* [2004] as in *Filipiak et al.* [2012], for the "d" parameter. From Equation 2, this parameter only contributes during the cooling part of the model. The averaged monthly diurnal cycle from April to July showed that no change occurs from using the climatological mixed layer depth instead of the actual bathymetry, at least for the Northern European Seas.

The FMKLB scheme is based on the statistical distribution of warming, but has been developed using SEVIRI data over different regions from the focus area of this study. Attempts to modify the model and tune it to the more local conditions of the North and the Baltic seas were made by examining the possibility of including the mean attenuation coefficient at 490 nm as an additional parameter. Results from regression analyses (not shown here) indicate that $K_{d(490)}$ does contribute to the total variance of the dSST signal from the FMKLB model. Unfortunately, the significance of this contribution is rather low. A long period of coincident modeled anomalies and $K_{d(490)}$ values could possibly strengthen this signal and this is something to look at in the future. In addition, the $K_{d(490)}$ values are daily and this is a limitation when attempting to correlate it with the hourly model results, as any variability in the $K_{d(490)}$ occurs in longer time scales.

In *Zeng & Beljaars* [2005], the shape exponent of the warm layer ν was set to a value of 0.3 for a warm layer depth of d =3 m. They also state that if the depth d is significantly altered, the ν parameter has to be changed. In this study, the combination of parameters described in *Zeng & Beljaars* [2005] was found to over predict warming in most areas. This may be related to a potential low bias of the HIRLAM winds in these areas and to the fact that different input fields were used for the foundation temperature. In addition, the ν parameter has not been altered for d_2 =6 m. Future work on the sensitivity of this parameter in combination with the depth d of the diurnal-free layer may provide a more appropriate combination of the two, which nonetheless may be very local. Sensitivity tests were performed for point locations of the buoys and platforms of the MARNET network in the North Sea and the Baltic Sea (see *Karagali et al.* [2012], Figure 1b). The depth of the warm layer d was allowed to vary between 2 m and 6 m and the shape exponent ν varied between 0.1 and 0.9. The results of the ZB model with various configurations were compared to the SEVIRI observed warming at these point locations; the collocation criteria were set such that the grid cell containing each location was selected. For the test period 03/01/2009 to 12/31/2009 it was found that most SEVIRI–ZB biases were minimized (0-0.2 K) for the d parameter equal to 6 m and the ν parameter between 0.1-0.3. Most standard deviations were lowest (0.2-0.9 K) for ν ranging between 0.2 and 0.6. Lowest rmse values (0.24-0.94 K) were found for different ν depending on the location, but mostly ranged between 0.2 and 0.5. The correlation was maximized (0.96-0.98) for almost all locations, for the combination d =6 m and ν =0.1 and the number of match-ups ranged from 1000 to 1600.

Correlation between SEVIRI and U, Q, SSI, SSI' and $K_{d(490)}$ showed the correct sign but weak signals. Especially for the HIRLAM fields, such findings are positive especially under the assumption that the model can not exactly resolve the spatial location of low winds where diurnal warming is observed. In addition, the correlation between Q and

SSI' is very high, on the order of 0.9 which provides a positive validation of the HIRLAM fluxes and the computation of Q. A higher correlation of the SEVIRI dSST with $K_{d(490)}$ was expected. As the temporal variability of $K_{d(490)}$ is expected to be lower than that of SST, it is rather hard to associate hourly variations of temperature with variations in $K_{d(490)}$. Despite that, investigations on the dependence of SEVIRI dSST$\geq$2 K on Q and U from HIRLAM for different ranges of $K_{d(490)}$, showed that for high $K_{d(490)}$ values (>0.37 m^{-1}, i.e. the upper 25th percentile) dSST was higher for very low wind speeds compared to the cases when $K_{d(490)}$ was lower than 0.37 m^{-1}. Such findings highlight the contribution of $K_{d(490)}$ to increased dSST values, given sufficiently low wind and high insolation values.

The modeled dSST negative correlations with $K_{d(490)}$ are somewhat difficult to interpret. It may be due to the time-lag between the observed and modeled dSST as the selection is controlled by the SEVIRI anomalies. This assumption is supported by the fact that the highest (negative) correlation is found for the ZB 6 m version, which also shows the greatest time lag. Independent sensitivity analyses of the dependence of modeled dSST on Q and U from HIRLAM for different ranges of $K_{d(490)}$ were performed similar to ones performed for SEVIRI dSSTs. All schemes showed similar behavior for $K_{d(490)}$ higher than 0.37 m^{-1} as for SEVIRI dSSTs. Especially the ZB d_1 version, showed very appropriate stratification and a significant decrease of dSST with increasing U.

6. Conclusions

The aim of this study is to evaluate diurnal warming in the Northern European Seas through simple parameterizations using as input fields the outputs of a regional NWP model and to compare the performance of the different parameterizations with observed anomalies from SEVIRI. Another objective is to examine the correlation of the observed warming with different physical parameters, such as wind speed U and integrated net heat flux Q from the regional NWP model, SSI from SEVIRI and the diffuse attenuation coefficient $K_{d(490)}$ that describes the optical properties of water.

Three different models have been applied, described in *Filipiak et al.* [2012], *Zeng & Beljaars* [2005] and *Clayson & Curry* [1996], due to their simple implementation and low computational cost. Of the three models applied in this study, only the first two resolve the full diurnal cycle and the third estimates peak skin warming. Collocated peak warming cases indicate almost zero biases for the *Zeng & Beljaars* [2005] and *Filipiak et al.* [2012] schemes. A 0.25 bias is found for the *Clayson & Curry* [1996] scheme, due to the difference between skin and sub-skin values. The overall correlation of peak warming between the models and SEVIRI does not exceed 0.45.

The distribution of dSSTs$\geq$2 K from the *Filipiak et al.* [2012] scheme simulates well that of the SEVIRI dSSTs, highlighting the ability of the model to capture the statistical distribution of warming. Both the *Filipiak et al.* [2012] and *Zeng & Beljaars* [2005] schemes identify more warming later in the day compared to SEVIRI, with the former scheme peaking 1 hour earlier and the latter peaking either at the same time or 1 hour later, depending on the version. Of the total SEVIRI dSSTs$\geq$2 K, 75% do not last more than 3 hours similar to the *Filipiak et al.* [2012] and *Zeng & Beljaars* [2005] d_2 version, while for the d_1 version this percentile is up to 6 hours.

The correlation between observed and modeled dSST and parameters such as U and Q from HIRLAM, SEVIRI SSI and $K_{d(490)}$ shows the correct trends. Negative correlation with U is found, low for SEVIRI and rather large for the models, as they are forced by the field. Positive correlation is found for Q, SSI and SSI', higher for the models (∼0.1-0.4) than for SEVIRI dSSTs (∼0-0.2). SEVIRI correlations with U and SSI are highest for the hours prior to the time of observed warming.

It is shown that the FMKLB model is representative of the distribution of warming cases as observed from SEVIRI. The ZB scheme is able to better resolve the spatial extent of the observed warming also accounting for the depth of the warm layer. This is an important parameter to be included in the models, especially for the Northern European Seas where water turbidity is high due to shallow depths, tidal currents and outflow of major rivers. The CC scheme provides a quick estimate of the peak skin warming which is useful as an estimation of potential warming but may not be accurate enough, as the "cool skin" bias can be significant.

Acknowledgments. This study was partially funded by the EU NORSEWInD project (TREN-FP7EN-219048) and the ESA STSE SSTDV: R.EX.–IM.A.M project. SEVIRI data were processed by the Centre de Météorolgie Spatiale, Métèo France and the OSI-SAF project. HIRLAM and $K_{d(490)}$ data obtained from the Danish Meteorological Institute (DMI), MARCOAST project. $K_{d(490)}$ data are also obtained from the GlobColour Project, through ftp. Meteorological mast data from Horns Rev were obtained through the NORSEWInD project and are courtesy of Vattenfall and DONG Energy. The authors would like to thank Chris Merchant for his help with the FMKLB model and his suggestions and guidance during the two months that Ioanna Karagali spent in the Department of Geosciences, University of Edinburgh. We are deeply grateful to Alec Bogdanoff and Carol Anne Clayson for their collaboration and for providing the Neural Network code and the code of their model. Finally, the authors would like to thank Peter Cornillon and an anonymous reviewer for their valuable comments and thorough revision of the manuscript.

References

Aminou, D.M.A., B. Jacquet, and F. Pasternak (1997). Characteristics of the Meteosat Second Generation (MSG) radiometer/imager: SEVIRI. *Proc. SPIE 3221, 19*, doi:10.1117/12.298084

Clayson, C.A., and J.A. Curry (1996). Determination of surface turbulent fluxes for the Tropical Ocean-Global Atmosphere Coupled Ocean-Atmosphere Response Experiment: comparison of satellite retrievals and in situ measurements. *J. Geophys. Res., 101(C12)*, 28515–28528.

Clayson, C.A., and D. Weitlich (2007). Variability of tropical diurnal sea surface temperature. *J. Climate, 20(2)*, 334–352.

de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res., 109*, C12003, doi:10.1029/2004JC002378.

Fairall, C.W., E.F. Bradley, J.S. Godfrey, G.A. Wick, J.B. Edson, and G.S. Young (1996a). Cool-skin and warm-layer effects on sea surface temperature, *J. Geophys. Res., 101(C1)*, 1295–1308, doi:10.1029/95JC03190

Filipiak, M.J., C.J. Merchant, H. Kettle, and P. Le Borgne (2010). A statistical model for sea surface diurnal warming driven by numerical weather prediction fluxes and winds. *Ocean Sci., 7(4)*, 197–209.

Gentemann, C.L., C.J. Donlon, A. Stuart-Menteth, and F.J. Wentz (2003). Diurnal signals in satellite sea surface temperature measurements. *Geophys. Res. Lett., 30(3)*, 40, doi:10.1029/2002GL016291.

Gentemann, C.L., P.J. Minnett, and B. Ward (2009). Profiles of Ocean Surface Heating (POSH): A new model of upper ocean diurnal warming. *J. Geophys. Res., 114(C7)*, C07017.

ESA GlobColour (2010). ESA DUE GlobColour Global Ocean Colour for Carbon Cycle Research Product User Guide. Version 1.4, GC-UM-ACR-PUG-01.

Kawai, Y., and H. Kawamura (2002). Evaluation of the diurnal warming of sea surface temperature using satellite-derived marine meteorological data. *J. Oceanography, 58(6)*, 805–814.

Kawai, Y., and A. Wada (2007). Diurnal sea surface temperature variation and its impact on the atmosphere and ocean: a review. *J. Oceanography, 63(5)*, 721–744.

Karagali, I., J.L. Høyer, and C.B. Hasager (2012). SST diurnal variability in the North Sea and the Baltic Sea. *Rem. Sens. Env., 121*, 159–170, 10.1016/j.rse.2012.01.016

Le Borgne, P., G. Legendre, and S. Péré (2012). Comparison of MSG/SEVIRI and drifting buoy derived diurnal warming estimates. *Rem. Sens. Env., 124*, 622–626, 10.1016/j.rse.2012.06.015

Mahura, A., K. Sattler, C. Petersen, B. Amstrup, and A. Baklanov (2005). DMI-HIRLAM Modelling with High Resolution Setup and Simulations for Areas of Denmark. *DMI Technical Report 05-12*, 44.

Merchant, C.J., M.J. Filipiak, P. Le Borgne, H. Roquet., E. Autret, et al. (2008). Diurnal warm-layer events in the western Mediterranean and European shelf seas. *Geophys. Res.*

Lett. 35(4), L04601.

Météo France (2005). Surface Solar Irradiance Product Manual, Ocean & Sea Ice SAF, Version 1.5, SAF/OSI/M-F/TEC/MA/123, November 2005.

Peters, S.W.M., M. Eleveld, R. Pasterkamp, H. Van der Woerd, M. Devolder, et al. (2005). Atlas of Chlorophyll-a concentration for the North Sea based on MERIS imagery of 2003. *Vrije Universiteit, Amsterdam, edition 3.0*

Price, J.F., R.A. Weller, and R. Pinkel (1986). Diurnal Cycling: Observations and Models of the Upper Ocean Response to Diurnal Heating, Cooling and Wind Mixing. *J. Geophys. Res., 91(C7)*, 8411–8427.

Webster, P.J., C.A. Clayson, and J.A. Curry (1996). Clouds, radiation, and the diurnal cycle of sea surface temperature in the Tropical Western Pacific. *J. Clim., 9(8)*, 1712–1730.

Zeng, X., and A. Beljaars (2005). A prognostic scheme of sea surface skin temperature for modeling and data assimilation. *Geophys. Res. Lett., 32(14)*, L14605.

Ward, B. (2006). Near-surface ocean temperature. *J. Geophys. Res., 111(C2)*, C02004.

Table 1. Spatial correlation of dSST≥2 Kelvin between SEVIRI, FMKLB and the ZB schemes, from the spatial extend of warming shown in Figure 8 and with the addition of random noise.

Model	FMKLB		ZB d_1		ZB d_2	
	–	Noise	–	Noise	–	Noise
r	0.11	0.13	0.32	0.33	0.28	0.32

Table 2. Statistics of SEVIRI−modeled dSST$_{max}$. The bias, in Kelvin, is defined as SEVIRI−model dSST$_{max}$. σ is the standard deviation in Kelvin, RMSE the root mean squared error in K, r the correlation coefficient and N is the number of collocated cases.

Model	FMKLB	ZB d_1	ZB d_2	CC
Mean Bias	0.01	-0.03	-0.05	0.25
σ	0.59	0.76	0.64	0.59
RMSE	0.59	0.76	0.64	0.64
r	0.45	0.42	0.45	0.43
N	2613317	2613317	2613317	2613317

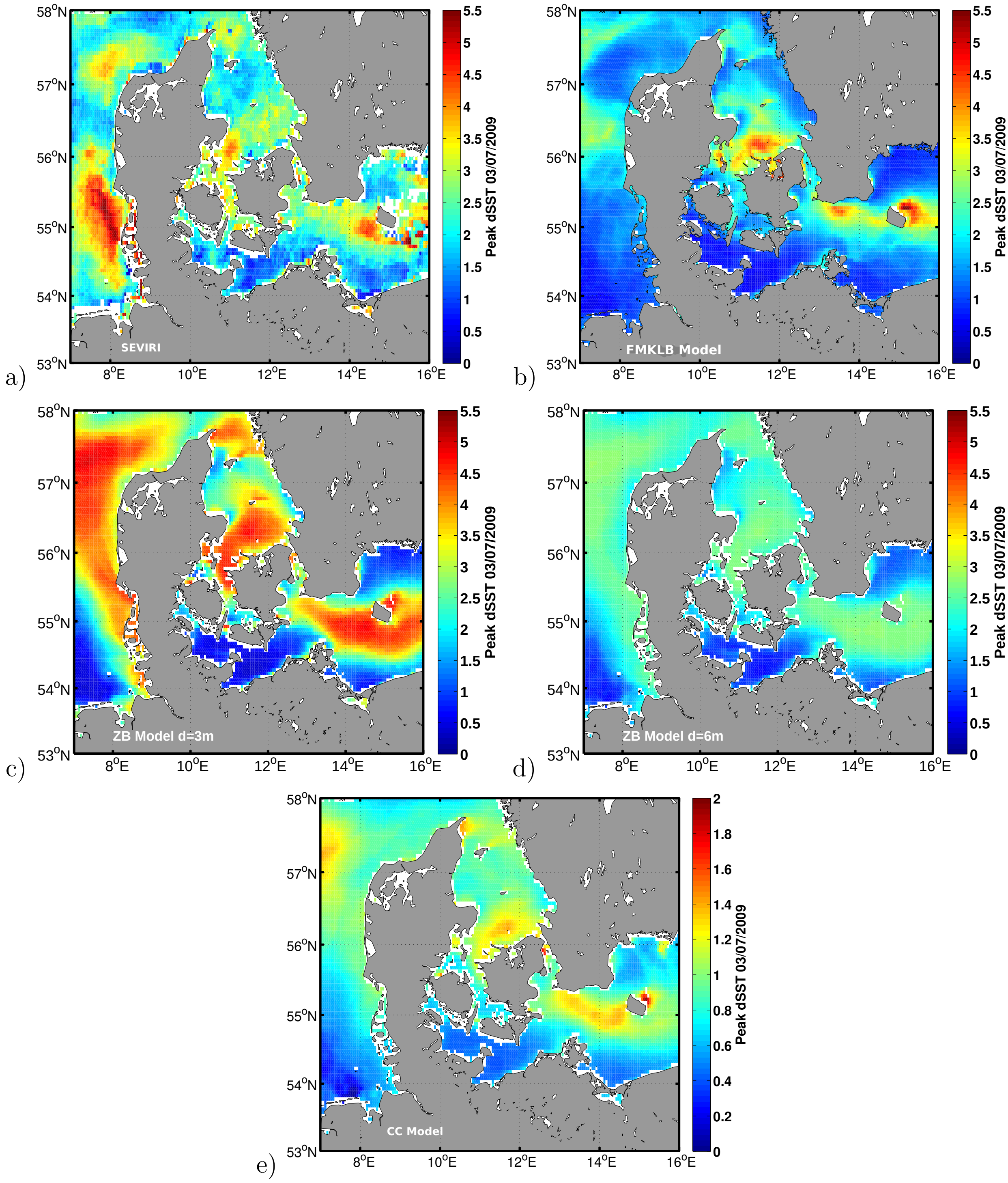

Figure 1. dSST$_{max}$ on the 03/07/2009, observed from a) SEVIRI and reproduced by the b) FMKLB, c-d) ZB and e) CC models. SEVIRI flagged with quality $\geq$3 are used to avoid gaps in the field. No random noise is added to the models. The CC dSST$_{max}$ has a reduced amplitude, to make the spatial variability visible.

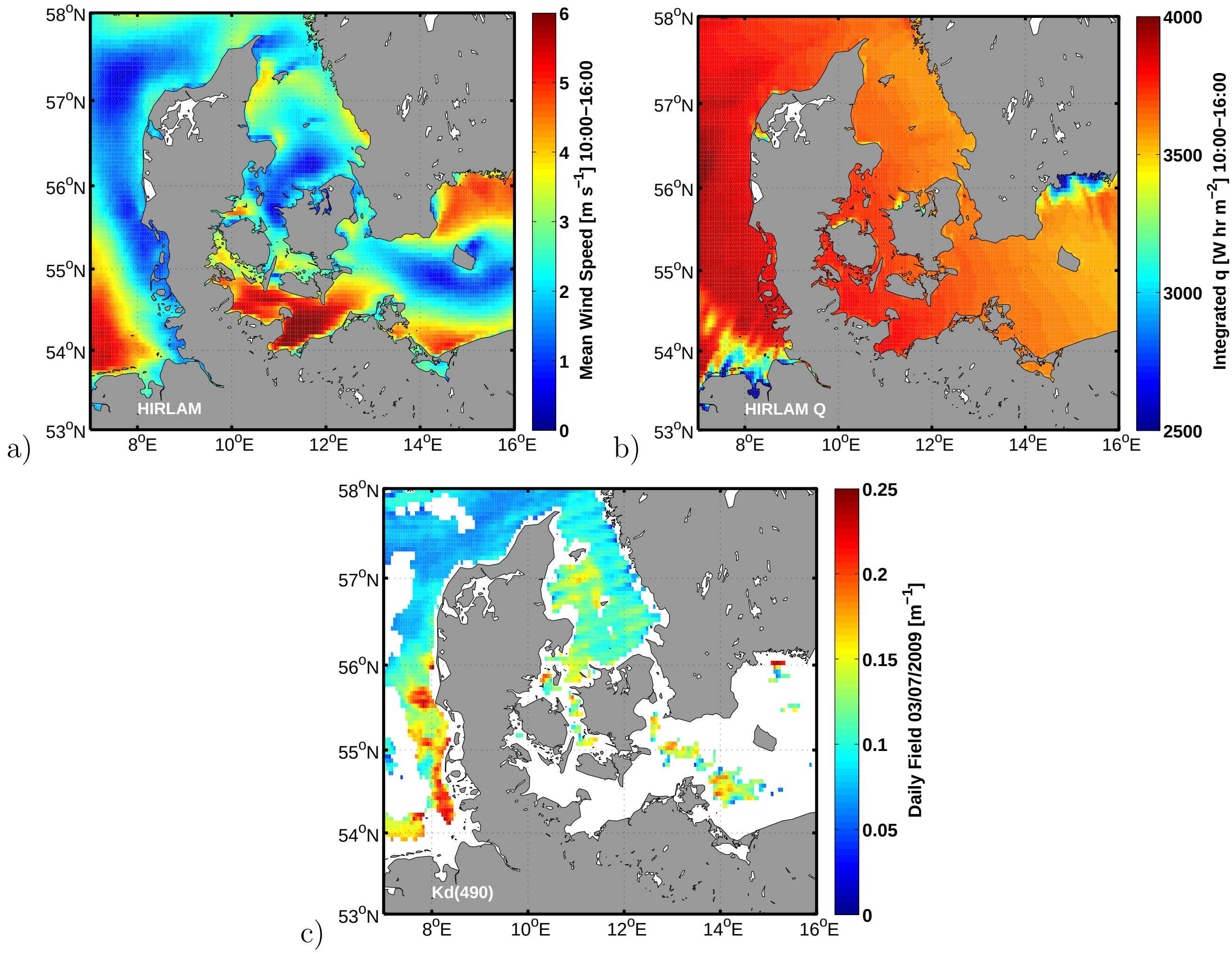

Figure 2. a) HIRLAM mean wind speed and b) integrated net heat flux from 10:00 to 16:00 and c) the daily $K_{d(490)}$ field from GlobColour for the diurnal warming event of 03/07/2009 shown in Figure 1.

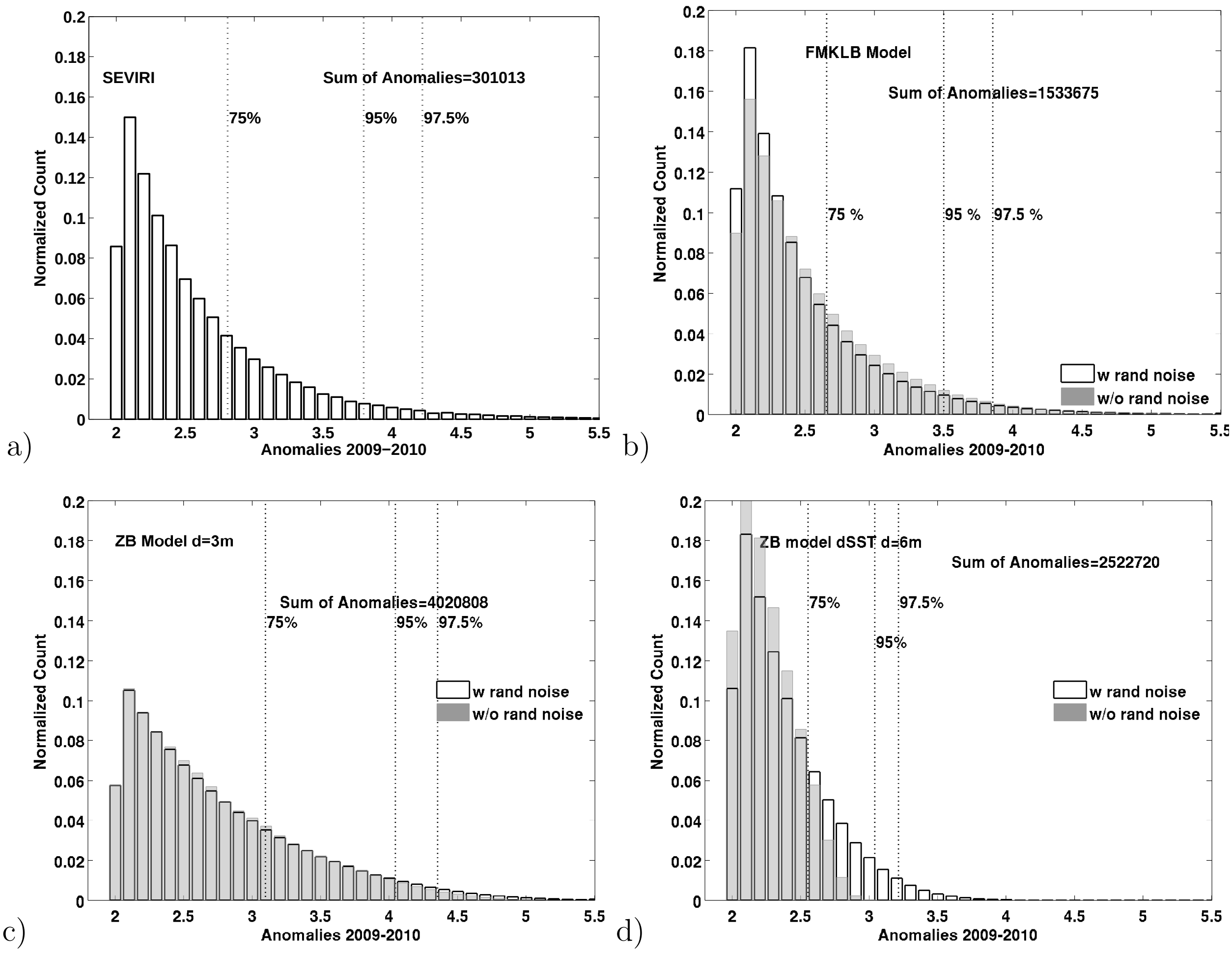

Figure 3. Distribution of anomalies≥2 K (02/2009–01/2010) from a) SEVIRI and the b) FMKLB and c-d) ZB schemes. The gray bars indicate the raw model outputs while the white bars outlined with black lines show the results when random noise has been added to the hourly model dSST.

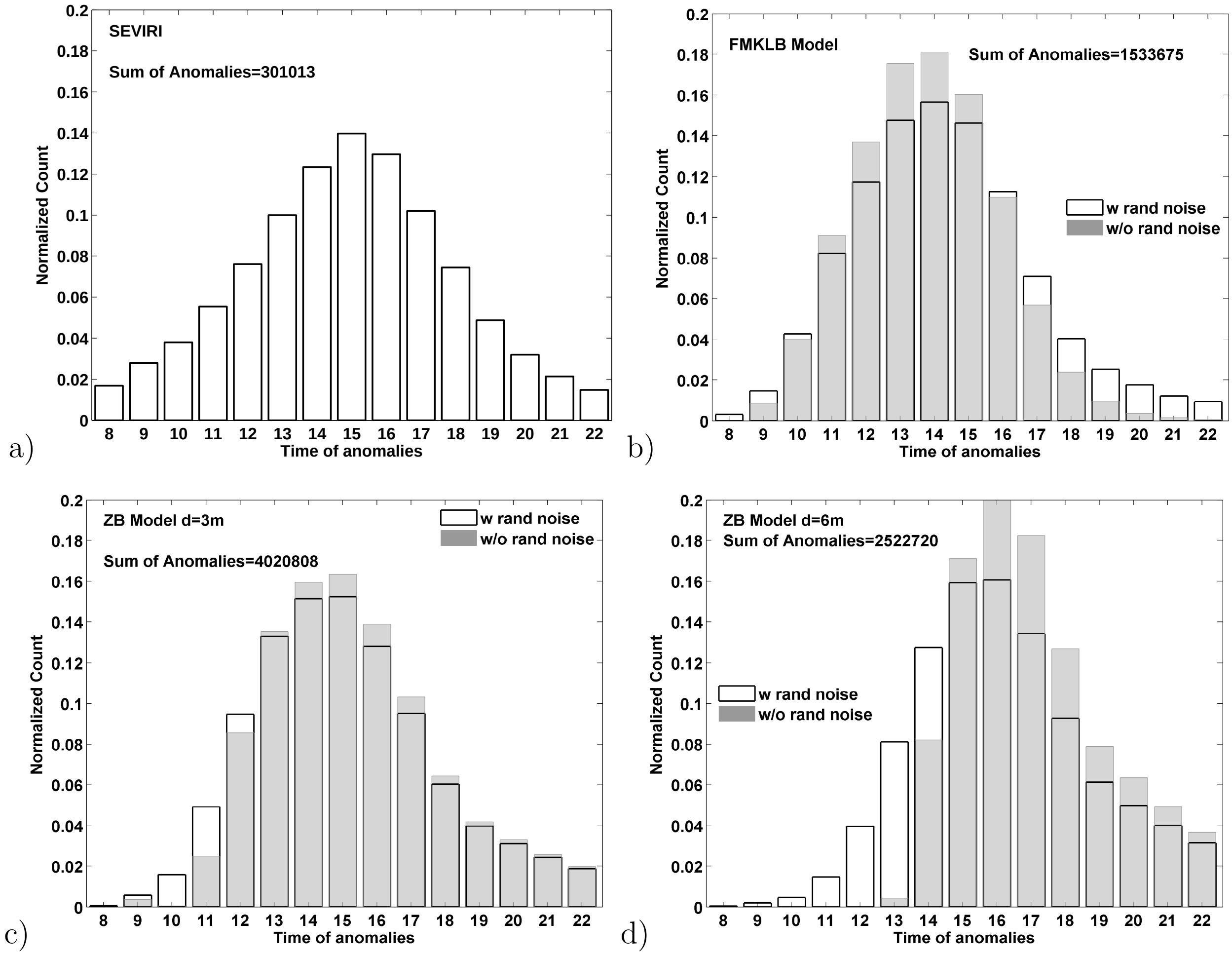

Figure 4. Distribution for the time of anomalies≥2 K (02/2009–01/2010) from a) SEVIRI and the b) FMKLB and c-d) ZB schemes. The gray bars indicate raw model outputs, white bars outlined with black lines show the noise-modified model dSST.

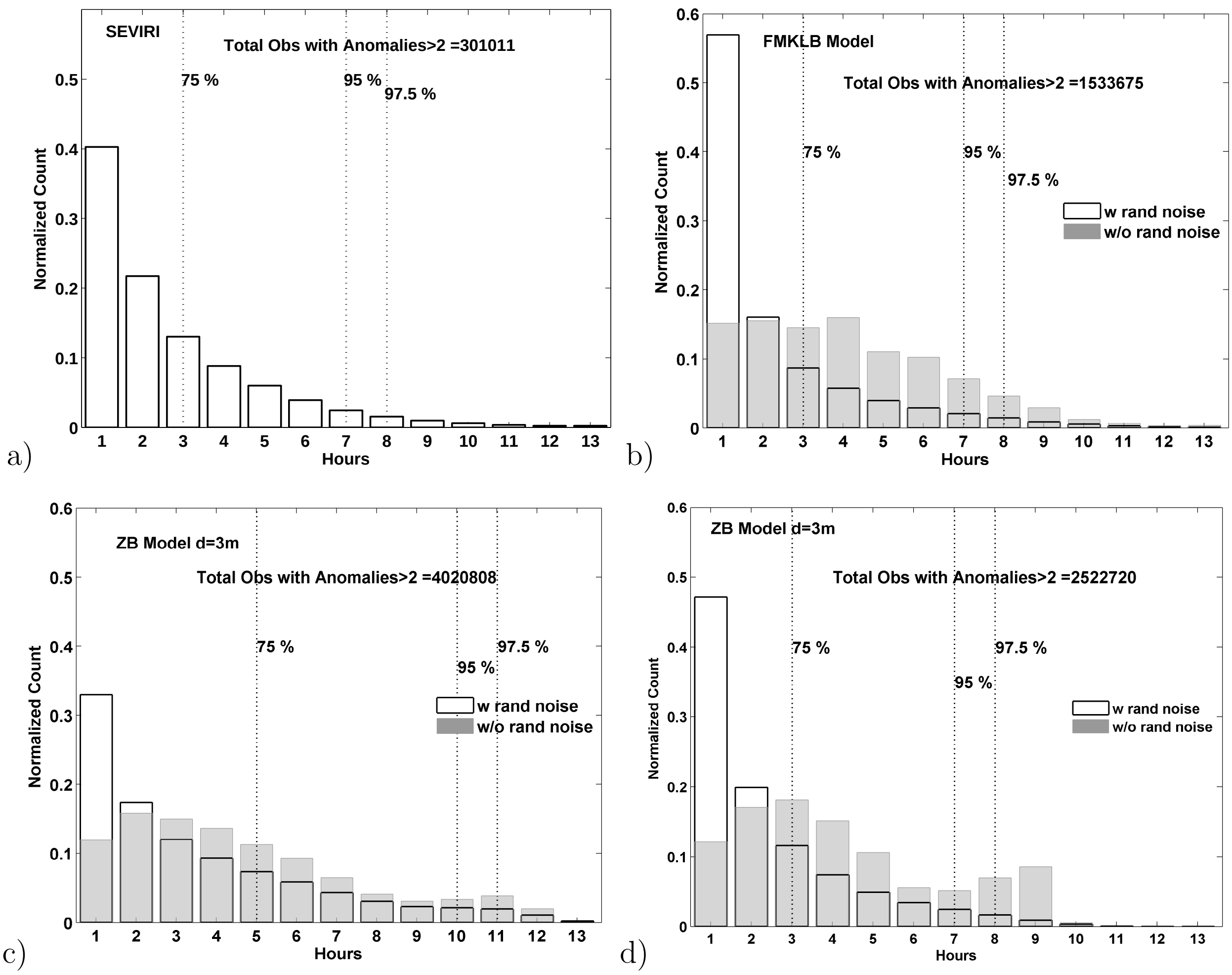

Figure 5. Duration of warming ≥2 K (02/2009–01/2010) from a) SEVIRI and the b) FMKLB and c-d) ZB schemes. Gray bars indicate the raw model outputs, white bars outlined with black lines show the noise-modified hourly model dSST. The dotted lines represent percentiles.

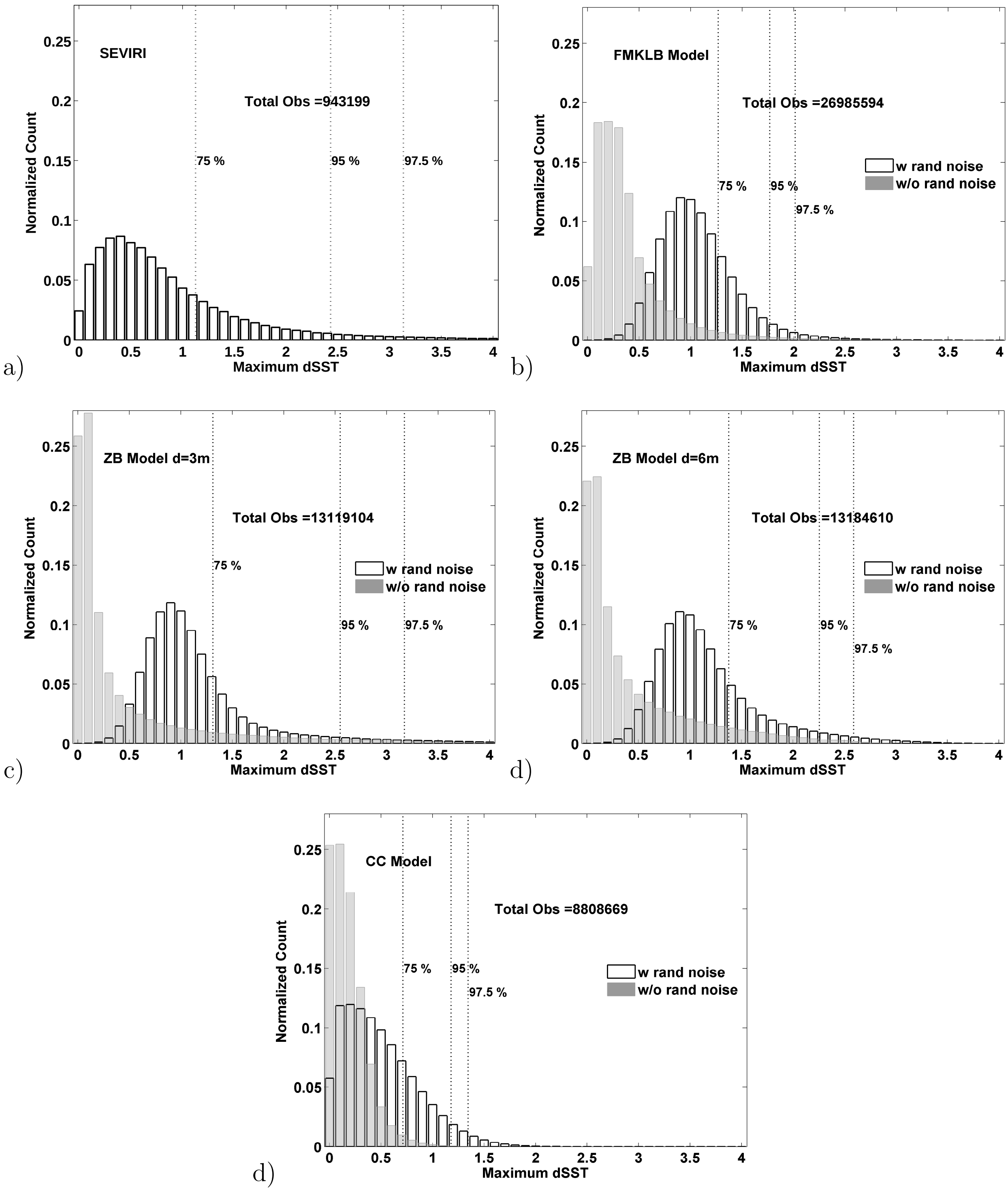

Figure 6. Distribution of maximum anomalies, counted for every grid cell (02/2009–01/2010), from a) SEVIRI and the b) FMKLB, c-d) ZB and e) CC schemes. Gray bars show raw model outputs, white bars outlined with black lines show the noise-modified hourly model dSST.

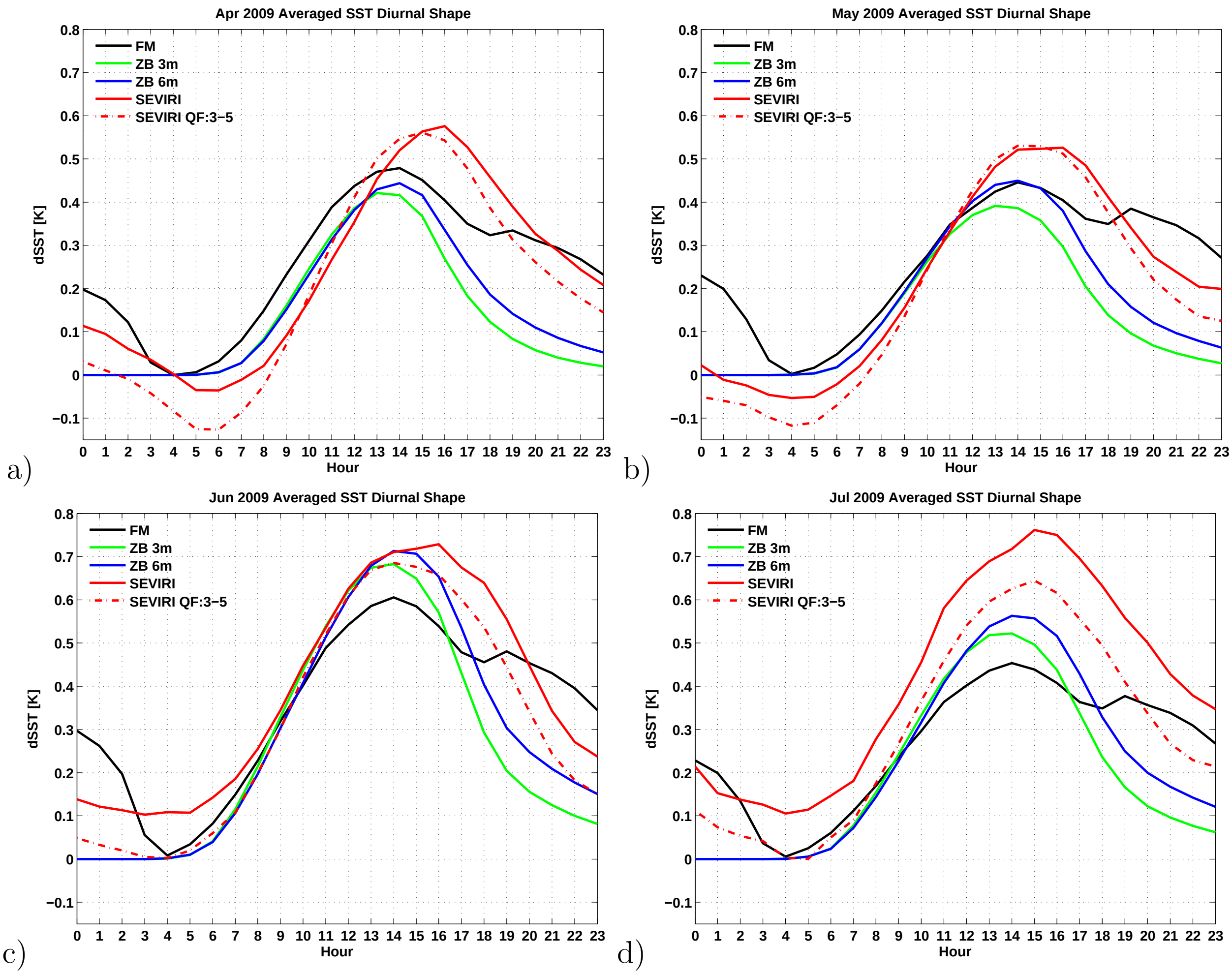

Figure 7. Monthly averaged diurnal cycle from a) SEVIRI (red), b) the FMKLB model (black), c) the ZB d_1 (green) and d) d_2 (blue). No anomaly threshold is applied, no time and space collocation criteria and no noise is added to the models.

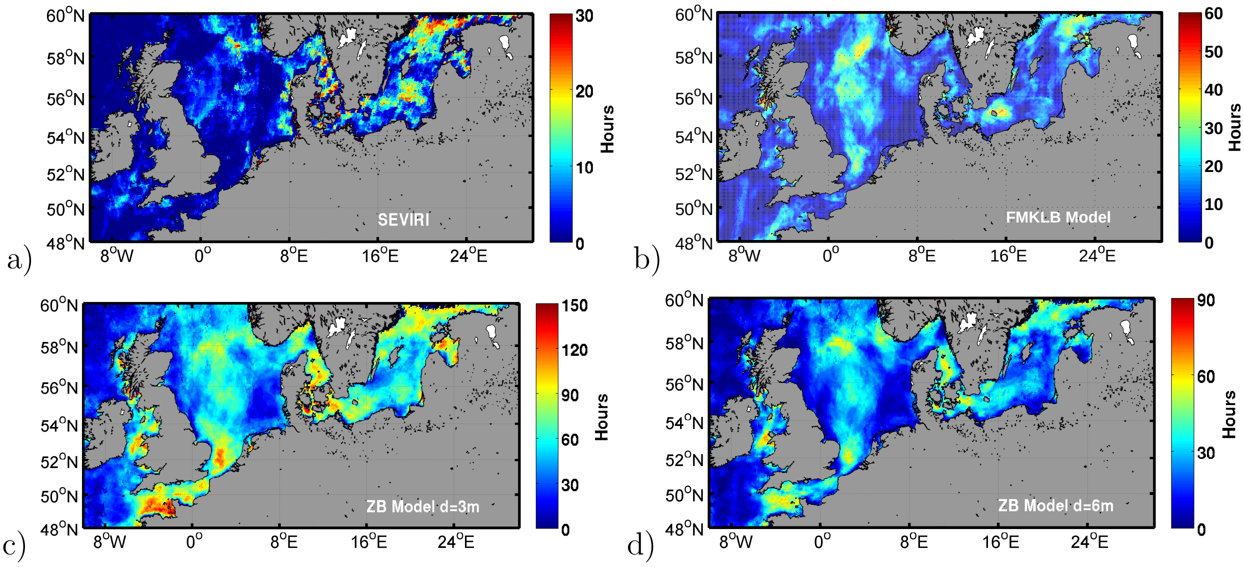

Figure 8. Spatial extent of anomalies≥2 K (02/2009–01/2010) from a) SEVIRI and the b) FMKLB, c-d) ZB schemes. The different scaling is to highlight the spatial variability. Maximum observed for SEVIRI is 30 hours, all others are multiples. No random noise added.

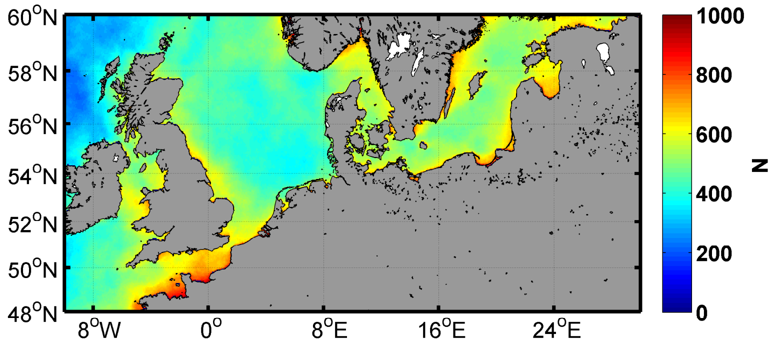

Figure 9. Number of occurrences for HIRLAM wind speed equal to or lower than 5 m s^{-1} and surface net heat flux higher than 300 W m^{-2}. From 02/2009 to 01/2010, a total of 8496 hourly fields are available.

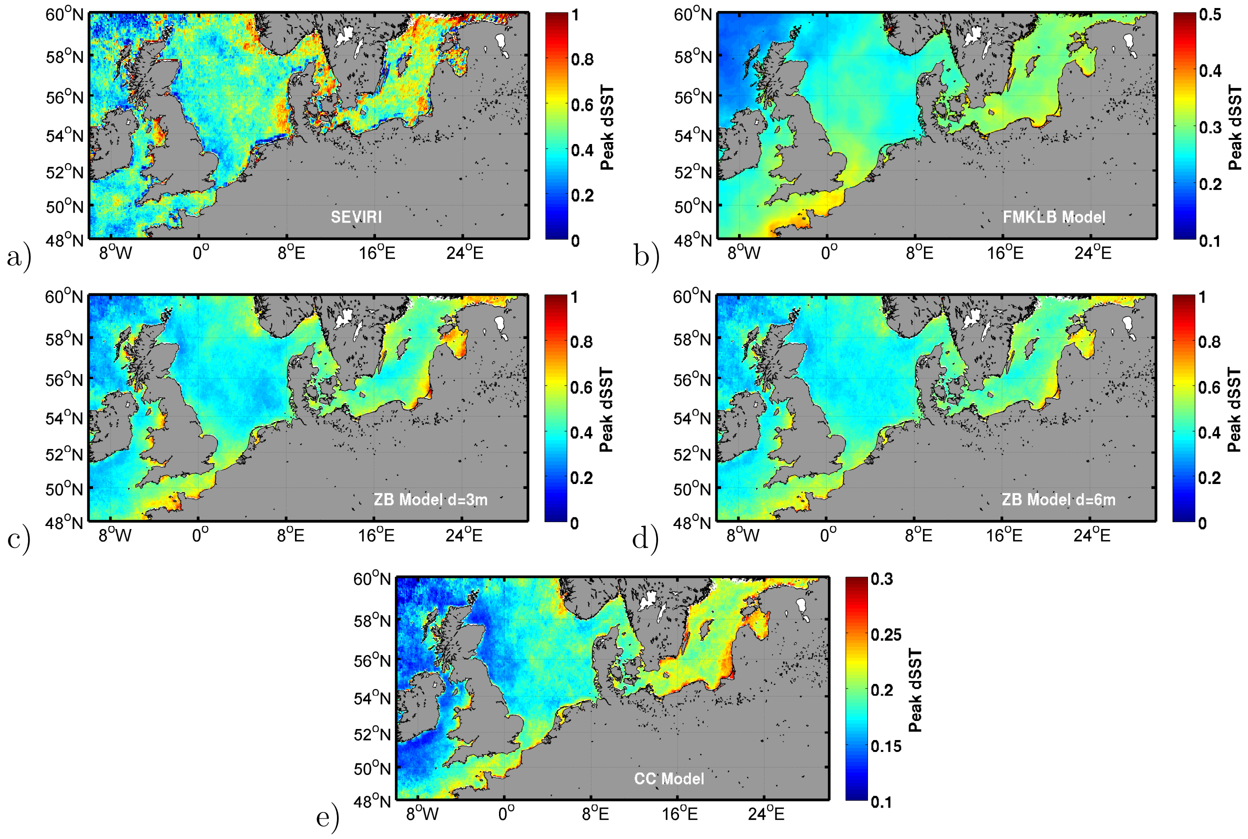

Figure 10. Spatial extent of averaged daily dSST$_{max}$ (02/2009–01/2010) from a) SEVIRI and the b) FMKLB, c-d) ZB and e) CC schemes. The reduced range of warming for the FMKLB and CC schemes is to highlight the spatial differences.

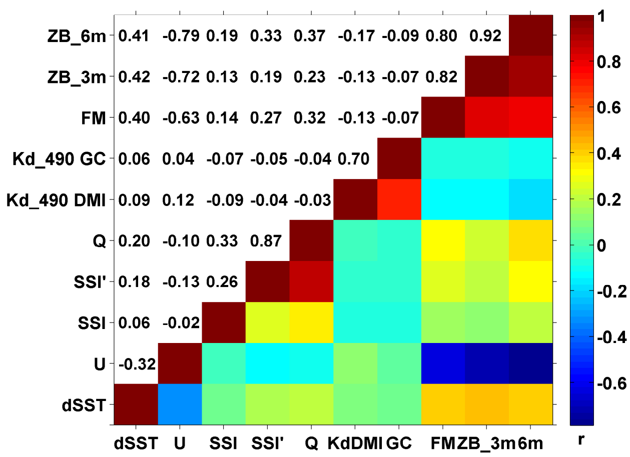

Figure 11. Correlation matrix of SEVIRI dSST, U, SSI, SSI′, Q and dSST from the FMKLB and ZB models (no random noise added).